



WWC Intervention Report

A summary of findings from a systematic review of the evidence



Primary Mathematics

Updated November 2015

Everyday Mathematics®

Program Description¹

Everyday Mathematics® is a core curriculum for students in grades pre-K–6. At each grade level, the curriculum provides students with multiple opportunities to reinforce concepts and practice skills. Across grade levels, concepts are reviewed and extended in varying instructional contexts. The distinguishing features of *Everyday Mathematics*® are its focus on real-life problem solving, student communication of mathematical thinking, and appropriate use of technology. The curriculum also emphasizes balancing different types of instruction (including collaborative learning), using various methods for skills practice, and fostering parent involvement in student learning.

Research²

The What Works Clearinghouse (WWC) identified one study of *Everyday Mathematics*® that both falls within the scope of the Primary Mathematics topic area and meets WWC group design standards. No studies meet WWC group design standards without reservations, and one study meets WWC group design standards with reservations. The study included 3,436 primary students in grades 3–5 in a large urban school district in Texas.

The WWC considers the extent of evidence for *Everyday Mathematics*® on the achievement outcomes of primary students to be small for mathematics achievement, the only outcome domain in Primary Mathematics. (See the Effectiveness Summary on p. 4 for more details of effectiveness by domain.)

Effectiveness

Everyday Mathematics® was found to have potentially positive effects on mathematics achievement for primary students.

Table 1. Summary of findings³

Outcome domain	Rating of effectiveness	Improvement index (percentile points)		Number of studies	Number of students	Extent of evidence
		Average	Range			
Mathematics achievement	Potentially positive effects	+11	na	1	3,436	Small

na = not applicable

Report Contents

Overview	p. 1
Program Information	p. 2
Research Summary	p. 3
Effectiveness Summary	p. 4
References	p. 5
Research Details for Each Study	p. 12
Outcome Measures for Each Domain	p. 14
Findings Included in the Rating for Each Outcome Domain	p. 15
Supplemental Findings for Each Outcome Domain	p. 16
Endnotes	p. 17
Rating Criteria	p. 18
Glossary of Terms	p. 19

This intervention report presents findings from a systematic review of *Everyday Mathematics*® conducted using the WWC Procedures and Standards Handbook, version 3.0, and the Primary Mathematics review protocol, version 3.1.

Program Information

Background

Everyday Mathematics® was developed by the University of Chicago School Mathematics Project and is published by McGraw-Hill Education. Address: P.O. Box 182605, Columbus, OH 43218. Email: customer.service@mheducation.com. Website: www.mheducation.com; www.everydaymath.com. Telephone: (800) 338-3987.

Program details

Everyday Mathematics® is a comprehensive curriculum for students in grades pre-K–6 that exposes students to material over extended periods of time by repeatedly reinforcing material both within and across grade levels. The curriculum includes six key features:

- **Real-life problem solving.** The curriculum emphasizes real-world applications of mathematics. Numbers, skills, and mathematics concepts are linked to typical real-life situations and contexts.
- **Balanced instruction.** Each lesson includes whole-class instruction and small-group, partner, or individual activities. This structure balances teacher-directed instruction with open-ended problem solving, hands-on explorations, long-term projects, and ongoing practice.
- **Multiple opportunities for basic skills practice.** The curriculum includes numerous opportunities for practicing basic skills, including written and oral fact practice, fluency activities and tests, mental math activities, daily sets of review problems, homework assignments, and a variety of math games.
- **Emphasis on communication.** Students are encouraged to explain and discuss their mathematical thinking, to clarify their thinking, and to gain insights from others.
- **Home/school partnerships.** Daily homework provides opportunities for family members to participate in students' mathematical learning. In addition, periodic letters are sent home to keep parents informed about their child's activities in mathematics.
- **Use of technology.** The curriculum includes many activities in which student learning is extended and enhanced through the use of calculators. Activities intended to reinforce paper and pencil and mental computation skills are clearly marked with an icon to indicate that calculator use is not permitted.

The curriculum is organized around Program Goals and Grade-Level Goals. Program Goals are organized by content strand and are the same at all grade levels. This structure interconnects the curriculum across grades and helps students build a solid base and consistently learn each skill and concept. Program Goals are refined through Grade-Level Goals, which specify the content that students are expected to master in each Program Goal for the year. There are approximately 20–25 Grade-Level Goals for each grade.

Curriculum materials include teacher guides with lesson plans and family letters, student materials, manipulatives, games, and online resources.

Cost

Curriculum sets for *Everyday Mathematics*® are bundled by grade and are available for pre-K–6. The pre-K and kindergarten classroom resource sets are \$167.76 and \$206.13, respectively. The classroom resource packages for grades 1–6 cost \$263.52. Additional materials range in cost from \$8.58 for a Skills Link Student Book to \$526.65 for a classroom Manipulative Kit.

Research Summary

The WWC identified 34 eligible studies that investigated the effects of *Everyday Mathematics*® on mathematics achievement for primary students. An additional 58 studies were identified but do not meet WWC eligibility criteria for review in this topic area. Citations for all 92 studies are in the References section, which begins on p. 5.

The WWC reviewed 34 eligible studies against group design standards. One study (Waite, 2000) uses a quasi-experimental design that meets WWC group design standards with reservations. The study is summarized in this report. Thirty-three studies do not meet WWC group design standards.

Table 2. Scope of reviewed research

Grades	3–5
Delivery method	Whole class
Program type	Curriculum

Summary of studies meeting WWC group design standards without reservations

No studies of *Everyday Mathematics*® met WWC group design standards without reservations.

Summary of study meeting WWC group design standards with reservations

Waite (2000) examined the effect of *Everyday Mathematics*® on the mathematics achievement of 732 third-, fourth-, and fifth-grade students in six schools against a comparison group of 2,704 third-, fourth-, and fifth-grade students in 12 similar schools that were matched on baseline mathematics achievement scores, student demographics, and geographical location. The schools in the intervention group were in their first year of implementing the first version of *Everyday Mathematics*®. The comparison group used a traditional mathematics curriculum approved by the school district.

Effectiveness Summary

The WWC review of *Everyday Mathematics*® for the Primary Mathematics topic area includes student outcomes in one domain: mathematics achievement. The findings below present the author’s estimates and WWC-calculated estimates of the size and statistical significance of the effects of *Everyday Mathematics*® on primary students. Additional comparisons are presented as supplemental findings in Appendix D. The supplemental findings do not factor into the intervention’s rating of effectiveness. For a more detailed description of the rating of effectiveness and extent of evidence criteria, see the WWC Rating Criteria on p. 18.

Summary of effectiveness for the mathematics achievement domain

One study that meets WWC group design standards with reservations reported findings in the mathematics achievement domain.

Waite (2000) reported a statistically significant positive effect of *Everyday Mathematics*® on mathematics achievement as measured by the total math score on the Texas Assessment of Academic Skills (TAAS) test. Based on WWC calculations, this effect was not statistically significant once corrections for clustering were made. However, the WWC determined that the effect size was substantively important (that is, 0.25 or greater). Based on this one study, the WWC categorized the effect of *Everyday Mathematics*® on mathematics achievement as potentially positive, with a small extent of evidence. Waite (2000) also reported results on subtests of the TAAS (Concepts, Operations, and Problem Solving), which are presented in the supplemental findings. These subtest analyses do not factor into the intervention’s rating of effectiveness.

Thus, for the mathematics achievement domain, one study showed substantively important positive effects. This results in a rating of potentially positive effects, with a small extent of evidence.

Table 3. Rating of effectiveness and extent of evidence for the mathematics achievement domain

Rating of effectiveness	Criteria met
Potentially positive effects <i>Evidence of a positive effect with no overriding contrary evidence.</i>	In the one study that reported findings, the estimated impact of the intervention on outcomes in the <i>mathematics achievement</i> domain was positive and substantively important.
Extent of evidence	Criteria met
Small	One study that included 3,436 students in 18 schools reported evidence of effectiveness in the <i>mathematics achievement</i> domain.

References

Studies that meet WWC group design standards without reservations

None.

Study that meets WWC group design standards with reservations

Waite, R. D. (2000). *A study of the effects of Everyday Mathematics on student achievement of third-, fourth-, and fifth-grade students in a large north Texas urban school district* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9992659)

Studies that do not meet WWC group design standards

- ARC Center. (2000a). *Everyday Mathematics: Glendale, CA*. In *The ARC Center's implementation stories from the field*. Retrieved from <http://www.comap.com> The study does not meet WWC group design standards because equivalence of the analytic intervention and comparison groups is necessary and not demonstrated.
- ARC Center. (2000b). *Everyday Mathematics: Portage, WI*. In *The ARC Center's implementation stories from the field*. Retrieved from <http://www.comap.com> The study does not meet WWC group design standards because equivalence of the analytic intervention and comparison groups is necessary and not demonstrated.
- Briars, D. J. (2004, July). *The Pittsburgh story: Successes and challenges in implementing standards-based mathematics programs*. Paper presented at the meeting of the UCSMP Everyday Mathematics Leadership Institute, Lisle, IL. The study does not meet WWC group design standards because equivalence of the analytic intervention and comparison groups is necessary and not demonstrated.
- Carroll, W. M. (1993). *Mathematical knowledge of kindergarten and first-grade students in Everyday Mathematics*. Chicago: University of Chicago School Mathematics Project. The study does not meet WWC group design standards because equivalence of the analytic intervention and comparison groups is necessary and not demonstrated.
- Carroll, W. M. (1995). *Third-grade Everyday Mathematics students' performance on the 1993 and 1994 Illinois state mathematics test*. Chicago: University of Chicago School Mathematics Project. The study does not meet WWC group design standards because equivalence of the analytic intervention and comparison groups is necessary and not demonstrated.
- Carroll, W. M. (1996a). *A follow-up to the fifth-grade field test of Everyday Mathematics: Geometry and mental and written computation*. Chicago: University of Chicago School Mathematics Project, Elementary Component. The study does not meet WWC group design standards because equivalence of the analytic intervention and comparison groups is necessary and not demonstrated.
- Carroll, W. M. (1996b). Mental computation of students in a reform-based mathematics curriculum. *School Science and Mathematics*, 96(6), 305–311. The study does not meet WWC group design standards because equivalence of the analytic intervention and comparison groups is necessary and not demonstrated.
- Carroll, W. M. (1997). Results of third-grade students in a reform curriculum on the Illinois state mathematics test. *Journal for Research in Mathematics Education*, 28(2), 237–242. The study does not meet WWC group design standards because equivalence of the analytic intervention and comparison groups is necessary and not demonstrated.
- Carroll, W. M. (1998). Geometric knowledge of middle school students in a reform-based mathematics curriculum. *School Science and Mathematics*, 98(4), 188–197. The study does not meet WWC group design standards because equivalence of the analytic intervention and comparison groups is necessary and not demonstrated.
- Carroll, W. M. (2000). Invented computational procedures of students in a standards-based curriculum. *Journal of Mathematical Behavior*, 18(2), 111–121. The study does not meet WWC group design standards because equivalence of the analytic intervention and comparison groups is necessary and not demonstrated.

- Carroll, W. M. (2001). Students in a standards-based mathematics curriculum: Performance on the 1999 Illinois state achievement test. *Illinois Mathematics Teacher*, 52(1), 3–7. The study does not meet WWC group design standards because equivalence of the analytic intervention and comparison groups is necessary and not demonstrated.
- Carroll, W. M., & Porter, D. (1994). *A field test of fourth grade Everyday Mathematics: Summary report*. Chicago: University of Chicago School Mathematics Project, Elementary Component. The study does not meet WWC group design standards because equivalence of the analytic intervention and comparison groups is necessary and not demonstrated.

Additional source:

- Everyday Learning Corporation. (1996). *Everyday Mathematics student achievement studies (study: UCSMP fourth grade field test)*. Chicago: Author.
- Council of the Great City Schools. (2001). *Improving K-8 mathematics achievement in the Anchorage school district: Report of the strategic support team of the Council of the Great City Schools*. Washington, DC: Author. The study does not meet WWC group design standards because equivalence of the analytic intervention and comparison groups is necessary and not demonstrated.
- Cummins-Colburn, B. J. L. (2007). *Differences between state-adopted textbooks and student outcomes on the Texas Assessment of Knowledge and Skills examination* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3269605) The study does not meet WWC group design standards because equivalence of the analytic intervention and comparison groups is necessary and not demonstrated.
- Drueck, J. V., Fuson, K. C., Carroll, W. M., & Bell, M. S. (1995, April). *Performance of U.S. first graders in a reform math curriculum compared to Japanese, Chinese, and traditionally taught U.S. students*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco. The study does not meet WWC group design standards because equivalence of the analytic intervention and comparison groups is necessary and not demonstrated.
- Everyday Learning Corporation. (1996a). *Everyday Mathematics student achievement studies (study: Greensburg Salem and Everyday Mathematics)*. Chicago: Author. The study does not meet WWC group design standards because equivalence of the analytic intervention and comparison groups is necessary and not demonstrated.
- Everyday Learning Corporation. (1996b). *Everyday Mathematics student achievement studies (study: Illinois goals assessment program performance)*. Chicago: Author. The study does not meet WWC group design standards because equivalence of the analytic intervention and comparison groups is necessary and not demonstrated.
- Everyday Learning Corporation. (1996c). *Everyday Mathematics student achievement studies (study: Northwestern University analysis of students)*. Chicago: Author. The study does not meet WWC group design standards because equivalence of the analytic intervention and comparison groups is necessary and not demonstrated.
- Fuson, K. C., & Carroll, W. M. (n.d.). *Summary of comparison of Everyday Mathematics (EM) and McMillan (MC): Evanston student performance on whole-class tests in grades 1, 2, 3, and 4*. Unpublished manuscript. The study does not meet WWC group design standards because equivalence of the analytic intervention and comparison groups is necessary and not demonstrated.

Additional source:

- Carroll, W. M., & Fuson, K. C. (1998). *A comparison of Everyday Mathematics (EM) and McMillan (MC) on Evanston student performance on whole-class tests: Recommendations for revision of Everyday Mathematics grades 1, 2, 3, and 4*. Unpublished manuscript.
- Fuson, K. C., Carroll, W. M., & Drueck, J. V. (2000). Achievement results for second and third graders using the standards-based curriculum Everyday Mathematics. *Journal for Research in Mathematics Education*, 31(3), 277–295. The study does not meet WWC group design standards because equivalence of the analytic intervention and comparison groups is necessary and not demonstrated.

- Grady, M., Watkins, S., & Montalvo, G. (2012). The effect of constructivist mathematics on achievement in rural schools. *Rural Educator*, 33(3), 38–47. The study does not meet WWC group design standards because the measures of effectiveness cannot be attributed solely to the intervention.
- James, O. (2011). *A study of selected test-taking strategies and demographic factors on the math performance of at-risk elementary students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3536908) The study does not meet WWC group design standards because equivalence of the analytic intervention and comparison groups is necessary and not demonstrated.
- Mathematics Evaluation Committee of the Hopewell Valley Regional School District. (1997). *Mathematics evaluation report: Year two*. Pennington, NJ: Author. The study does not meet WWC group design standards because equivalence of the analytic intervention and comparison groups is necessary and not demonstrated.
- McCabe, K. J. (2001). Mathematics in our schools: An effort to improve mathematics literacy. *Masters Abstracts International*, 40(04), 835. The study does not meet WWC group design standards because equivalence of the analytic intervention and comparison groups is necessary and not demonstrated.
- Powell, T. L. (2014). *A comparative analysis of the Singapore Math curriculum and the Everyday Mathematics curriculum on fifth grade achievement in a large northeastern urban public school district* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 1564756065) The study does not meet WWC group design standards because equivalence of the analytic intervention and comparison groups is necessary and not demonstrated.
- Riordan, J. E., & Noyce, P. E. (2001). The impact of two standards-based mathematics curricula on student achievement in Massachusetts. *Journal for Research in Mathematics Education*, 32(4), 368–398. The study does not meet WWC group design standards because equivalence of the analytic intervention and comparison groups is necessary and not demonstrated.
- Roan, C. (2012). *A comparison of elementary mathematics achievement in Everyday Math and Saxon Math schools in Illinois* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3507509) The study does not meet WWC group design standards because equivalence of the analytic intervention and comparison groups is necessary and not demonstrated.
- Sconiers, S., Isaacs, A., Higgins, T., McBride, J., & Kelso, C. R. (2003). *The ARC Center tri-state student achievement study*. Lexington, MA: The Consortium for Mathematics and Its Applications. The study does not meet WWC group design standards because equivalence of the analytic intervention and comparison groups is necessary and not demonstrated.
- Additional source:**
- SRA/McGraw-Hill. (2001). *Everyday Mathematics student achievement studies: Volume 4 (study: Illinois, Massachusetts, and Washington achievement study)*. Chicago: Author.
- SRA/McGraw-Hill. (2001a). *Everyday Mathematics student achievement studies: Volume 3 (study: Kentucky Commonwealth Accountability Testing System)*. Chicago: Author. The study does not meet WWC group design standards because equivalence of the analytic intervention and comparison groups is necessary and not demonstrated.
- SRA/McGraw-Hill. (2001b). *Everyday Mathematics student achievement studies: Volume 3 (study: Massachusetts Comprehensive Assessment System)*. Chicago: Author. The study does not meet WWC group design standards because equivalence of the analytic intervention and comparison groups is necessary and not demonstrated.
- SRA/McGraw-Hill. (2001c). *Everyday Mathematics student achievement studies: Volume 3 (study: Pennsylvania State Assessment System)*. Chicago: Author. The study does not meet WWC group design standards because equivalence of the analytic intervention and comparison groups is necessary and not demonstrated.
- Ward, W. M. (2009). *The influence of a reform-based mathematics program on third, fourth, and fifth grade student achievement* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3440870) The study does not meet WWC group design standards because equivalence of the analytic intervention and comparison groups is necessary and not demonstrated.

Woodward, J., & Baxter, J. (1997). The effects of an innovative approach to mathematics on academically low-achieving students in inclusive settings. *Exceptional Children*, 63(3), 373–388. The study does not meet WWC group design standards because the measures of effectiveness cannot be attributed solely to the intervention.

Studies that are ineligible for review using the Primary Mathematics Evidence Review Protocol

- Allen, C. (2007). *An action based research study on how using manipulatives will increase students' achievement in mathematics*. Unpublished research study. Detroit, MI: Marygrove College. The study is ineligible for review because it does not use an eligible design.
- ARC Center. (2000). *Everyday Mathematics: Kent, WA*. In *The ARC Center's implementation stories from the field*. Retrieved from <http://www.comap.com> The study is ineligible for review because it does not use an eligible design.
- Baxter, J. A., Woodward, J., & Olson, D. (2001). Effects of reform-based mathematics instruction on low achievers in five third-grade classrooms. *Elementary School Journal*, 101(5), 529–547. The study is ineligible for review because it does not use an eligible design.
- Bhatt, R., Koedel, C., & Lehmann, D. (2013). Is curriculum quality uniform? Evidence from Florida. *Economics of Education Review*, 34, 107–121. The study is ineligible for review because it does not use an eligible design.
- Bieda, K. N., Ji, X., Drwencke, J., & Picard, A. (2014). Reasoning-and-proving opportunities in elementary mathematics textbooks. *International Journal of Educational Research*, 64, 71–80. The study is ineligible for review because it does not use an eligible design.
- Brehe Pixler, P. (2009). *Mild disability students and Everyday Mathematics: Serving the needs of this population* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3399155) The study is ineligible for review because it is out of scope of the protocol.
- Briars, D. J., & Resnick, L. B. (2000). *Standards, assessments—and what else? The essential elements of standards-based school improvement* (CSE Technical Report No. 528). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing. The study is ineligible for review because it does not use an eligible design.
- Carroll, W. M. (1995). *Report on the field test of fifth-grade Everyday Mathematics*. Chicago: University of Chicago School Mathematics Project, Elementary Component. The study is ineligible for review because it is out of scope of the protocol.
- Carroll, W. M. (1996). Use of invented algorithms by second graders in a reform mathematics curriculum. *Journal of Mathematical Behavior*, 15(2), 137–150. The study is ineligible for review because it does not use an eligible design.
- Carroll, W. M. (2001). *A longitudinal study of children in the Everyday Mathematics curriculum*. Retrieved from <http://everydaymath.uchicago.edu/> The study is ineligible for review because it does not use an eligible design.

Additional sources:

- Carroll, W. M., & Fuson, K. C. (1999). *Achievement results for fourth graders using the standards-based curriculum Everyday Mathematics*. Unpublished manuscript.
- Carroll, W. M., Fuson, K. C., & Drucek, J. D. (n.d.). *A longitudinal study of second and third graders using the reform curriculum Everyday Mathematics by the University of Chicago School Mathematics Project*. Unpublished manuscript.
- Carroll, W. M., Fuson, K. C., & Diamond, A. (2000). Use of student-constructed number stories in a reform-based curriculum. *Journal of Mathematical Behavior*, 19(1), 49–62. This study is ineligible for review because it does not use an eligible design.
- Carroll, W. M., & Isaacs, A. (2003). Achievement of students using the University of Chicago School Mathematics Project's Everyday Mathematics. In S. L. Senk & D. R. Thompson (Eds.), *Standards-based school mathematics curricula: What are they? What do students learn?* (pp. 79–108). Mahwah, NJ: Lawrence Erlbaum Associates, Inc. This study is ineligible for review because it does not use an eligible design.

- Chang, K., Males, L. M., Mosier, A., & Gonulates, F. (2011). Exploring U.S. textbooks' treatment of the estimation of linear measurements. *ZDM – International Journal on Mathematics Education*, 43(5), 697–708. This study is ineligible for review because it does not use an eligible design.
- Confrey, J. (2006). Comparing and contrasting the National Research Council report on evaluating curricular effectiveness with the What Works Clearinghouse approach. *Educational Evaluation and Policy Analysis*, 28(3), 195–213. doi:10.3102/01623737028003195 This study is ineligible for review because it does not use an eligible design.
- de Oliveira, L. C., & Cheng, D. (2011). Language and the multisemiotic nature of mathematics. *Reading Matrix: An International Online Journal*, 11(3), 255–268. This study is ineligible for review because it does not use an eligible design.
- Drueck, J. V. (1996, April). *Progression of multidigit addition and subtraction solution methods in high-, average-, and low-math-achieving second graders experiencing a reform curriculum*. Paper presented at the meeting of the American Educational Research Association, New York. This study is ineligible for review because it does not use an eligible design.
- Everyday Learning Corporation. (1996a). *Everyday Mathematics student achievement studies (study: Kalamazoo success story)*. Chicago: Author. This study is ineligible for review because it does not use an eligible design.
- Everyday Learning Corporation. (1996b). *Everyday Mathematics student achievement studies (study: UCSMP in Wheeling, Illinois)*. Chicago: Author. This study is ineligible for review because it does not use an eligible design.
- Feil, Y. C. (2010). *Can you teach in a normal way? Examining Chinese and U.S. curricula's approach to teaching fraction divisions* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3430859) The study is ineligible for review because it is out of scope of the protocol.
- Fraivillig, J. L. (1996). *Case studies and instructional frameworks of expert reform mathematics teaching* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9632689) This study is ineligible for review because it does not use an eligible design.
- Fraivillig, J. L., Murphy, L., & Fuson, K. C. (1999). Advancing children's mathematical thinking in Everyday Mathematics classrooms. *Journal for Research in Mathematics Education*, 30(2), 148–170. This study is ineligible for review because it does not use an eligible design.
- Additional sources:**
- Fraivillig, J. L. (2001). Strategies for advancing children's mathematical thinking. *Teaching Children Mathematics*, 7(8), 454–459.
- Fuson, K. (1997). What do we see in Everyday Mathematics classrooms? *Teacherlink*, 5, 1–2.
- Garelick, B. (2006). A tale of two countries and one school district. *Third Education Group Review/Essays*, 2(8). Retrieved from <http://www.thirddeducationgroup.org/> This study is ineligible for review because it does not use an eligible design.
- Garelick, B. (2006). A textbook case in textbook adoption. *Third Education Group Review*, 2(6). Retrieved from <http://nonpartisaneducation.org/> This study is ineligible for review because it does not use an eligible design.
- Hedges, L. V., & Stodolsky, S. S. (1988). *A follow-up of kindergarten Everyday Mathematics users* (Rep. No. 87/88-KEM-2). Chicago: University of Chicago School Mathematics Project. This study is ineligible for review because it does not use an eligible design.
- Hedges, L. V., Stodolsky, S. S., & Mathison, S. (1987). *A formative evaluation of kindergarten Everyday Mathematics* (Rep. No. 86/87-KEM-1). Chicago: University of Chicago School Mathematics Project. This study is ineligible for review because it does not use an eligible design.
- Helfant, M. T. (2005). *The relationship between third and fourth grade Everyday Mathematics assessments and performance on the New Jersey Assessment of Skills and Knowledge in fourth grade (NJASK/4)* (Unpublished doctoral dissertation). Seton Hall University, South Orange, NJ. This study is ineligible for review because it does not use an eligible design.

- Kiger, D., Herro, D., & Prunty, D. (2012). Examining the influence of a mobile learning intervention on third grade math achievement. *Journal of Research on Technology in Education*, 45(1), 61–82. This study is ineligible for review because it does not use an eligible design.
- Mac Iver, M. A., & Mac Iver, D. J. (2009). Urban middle-grade student mathematics achievement growth under comprehensive school reform. *Journal of Educational Research*, 102(3), 223–236. This study is ineligible for review because it is out of scope of the protocol.
- McCaffrey, D. F., Hamilton, L. S., Stecher, B. M., Klein, S. P., Bugliari, D., & Robyn, A. (2001). Interactions among instructional practices, curriculum, and student achievement: The case of standards-based high school mathematics. *Journal for Research in Mathematics Education*, 32(5), 493–517. This study is ineligible for review because it does not use a sample aligned with the protocol.
- Morton, M., & Montgomery, P. (2011). Youth empowerment programs for improving self-efficacy and self-esteem of adolescents. *Campbell Systematic Reviews*, 7(5). This study is ineligible for review because it is out of scope of the protocol.
- Murnane, R. J., & Nelson, R. R. (2005). *Improving the performance of the education sector: The valuable, challenging, and limited role of random assignment evaluations* (NBER Working Paper 11846). Cambridge, MA: National Bureau of Economic Research. This study is ineligible for review because it does not use an eligible design.
- Murphy, L. A. (1998). *Learning and affective issues among higher- and lower-achieving third-graders in math reform classrooms: Perspectives of children, parents, and teachers* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9913852) This study is ineligible for review because it does not use an eligible design.
- Northwestern Longitudinal Study of Everyday Mathematics. (1998). *Fourth-grade feedback on specific lessons*. Unpublished report. Evanston, IL: Northwestern University. This study is ineligible for review because it is out of scope of the protocol.
- Ramani, G. B., & Siegler, R. S. (2011). Reducing the gap in numerical knowledge between low- and middle-income preschoolers. *Journal of Applied Developmental Psychology*, 32(3), 146–159. This study is ineligible for review because it is out of scope of the protocol.
- Richards, A. L. (2010). *Improving the academic self-efficacy of middle school girls toward the study of mathematics through the use of theatrical infusion* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3428782) This study is ineligible for review because it is out of scope of the protocol.
- Saljo, R., & Wyndhamn, J. (1987). The formal setting as context for cognitive activities: An empirical study of arithmetic operations under conflicting premises for communication. *European Journal of Psychology of Education*, 2(3), 233–245. This study is ineligible for review because it is out of scope of the protocol.
- Saxe, G., Diakow, R., & Gearhart, M. (2013). Towards curricular coherence in integers and fractions: A study of the efficacy of a lesson sequence that uses the number line as the principal representational context. *ZDM—International Journal on Mathematics Education*, 45(3), 343–364. doi:10.1007/s11858-012-0466-2 This study is ineligible for review because it does not use an eligible design.
- Slavin, R. E. (2005). *Show me the evidence: Effective programs for elementary and secondary schools*. Unpublished manuscript. This study is ineligible for review because it does not use an eligible design.
- Slavin, R. E., & Lake, C. (2008). Effective programs in elementary mathematics: A best-evidence synthesis. *Review of Educational Research*, 78(3), 427–515. This study is ineligible for review because it does not use an eligible design.
- Son, J., & Senk, S. L. (2010). How reform curricula in the USA and Korea present multiplication and division of fractions. *Educational Studies in Mathematics*, 74(2), 117–142. This study is ineligible for review because it does not use an eligible design.
- Sood, S., & Jitendra, A. K. (2007). A comparative analysis of number sense instruction in reform-based and traditional mathematics textbooks. *Journal of Special Education*, 41(3), 145–157. This study is ineligible for review because it is out of scope of the protocol.

- SRA/McGraw-Hill. (2001a). *Everyday Mathematics student achievement studies: Volume 3 (study: California SAT-9)*. Chicago: Author. This study is ineligible for review because it does not use an eligible design.
- SRA/McGraw-Hill. (2001b). *Everyday Mathematics student achievement studies: Volume 3 (study: Florida Comprehensive Assessment Test)*. Chicago: Author. This study is ineligible for review because it does not use an eligible design.
- SRA/McGraw-Hill. (2001c). *Everyday Mathematics student achievement studies: Volume 3 (study: MAT-7 in Wichita, Kansas)*. Chicago: Author. This study is ineligible for review because it does not use an eligible design.
- SRA/McGraw-Hill. (2001d). *Everyday Mathematics student achievement studies: Volume 3 (study: Michigan Educational Assessment Program)*. Chicago: Author. This study is ineligible for review because it does not use an eligible design.
- SRA/McGraw-Hill. (2001e). *Everyday Mathematics student achievement studies: Volume 3 (study: SAT-9 in Santa Ana, California)*. Chicago: Author. This study is ineligible for review because it does not use an eligible design.
- SRA/McGraw-Hill. (2001f). *Everyday Mathematics student achievement studies: Volume 4 (study: Florida Comprehensive Assessment Test)*. Chicago: Author. This study is ineligible for review because it does not use an eligible design.
- SRA/McGraw-Hill. (2001g). *Everyday Mathematics student achievement studies: Volume 4 (study: Illinois Standards Achievement Test)*. Chicago: Author. This study is ineligible for review because it does not use an eligible design.
- SRA/McGraw-Hill. (2001h). *Everyday Mathematics student achievement studies: Volume 4 (study: Kentucky Core Content Test)*. Chicago: Author. This study is ineligible for review because it does not use an eligible design.
- SRA/McGraw-Hill. (2001i). *Everyday Mathematics student achievement studies: Volume 4 (study: Michigan Educational Assessment Program)*. Chicago: Author. This study is ineligible for review because it does not use an eligible design.
- SRA/McGraw-Hill. (2001j). *Everyday Mathematics student achievement studies: Volume 4 (study: North Carolina ABCs Accountability Model)*. Chicago: Author. This study is ineligible for review because it does not use an eligible design.
- SRA/McGraw-Hill. (2001k). *Everyday Mathematics student achievement studies: Volume 4 (study: South Carolina Palmetto Achievement Challenge Test)*. Chicago: Author. This study is ineligible for review because it does not use an eligible design.
- SRA/McGraw-Hill. (2001l). *Everyday Mathematics student achievement studies: Volume 4 (study: Stanford-9 in Tucson, Arizona)*. Chicago: Author. This study is ineligible for review because it does not use an eligible design.
- SRA/McGraw-Hill. (2001m). *Everyday Mathematics student achievement studies: Volume 4 (study: Tennessee Comprehensive Assessment Program)*. Chicago: Author. This study is ineligible for review because it does not use an eligible design.
- SRA/McGraw-Hill. (2001n). *Everyday Mathematics student achievement studies: Volume 4 (study: Texas Assessment of Academic Skills)*. Chicago: Author. This study is ineligible for review because it does not use an eligible design.
- SRA/McGraw-Hill. (2001o). *Everyday Mathematics student achievement studies: Volume 4 (study: Washington Assessment of Student Learning)*. Chicago: Author. This study is ineligible for review because it does not use an eligible design.
- Stein, M. K., & Kaufman, J. H. (2010). Selecting and supporting the use of mathematics curricula at scale. *American Educational Research Journal*, 47(3), 663–693. This study is ineligible for review because it is out of scope of the protocol.
- Van Dyke, F., & Tomback, J. (2005). Collaborating to introduce algebra. *Mathematics Teaching in the Middle School*, 10(5), 236–242. This study is ineligible for review because it does not use an eligible design.

Appendix A: Research details for Waite, 2000

Waite, R. D. (2000). *A study of the effects of Everyday Mathematics on student achievement of third-, fourth-, and fifth-grade students in a large north Texas urban school district* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9992659)

Table A. Summary of findings

Meets WWC group design standards with reservations

Outcome domain	Sample size	Study findings	
		Average improvement index (percentile points)	Statistically significant
Mathematics achievement	3,436 students	+ 11	No

Setting All schools in this study were located in a large urban school district in north Texas.

Study sample The study sample consisted of third-, fourth-, and fifth-grade students. Six schools within one district volunteered to implement the first edition of *Everyday Mathematics*® during the 1998–99 school year. A comparison group of 12 schools within the same school district was selected. Comparison schools did not use *Everyday Mathematics*® during the 1998–99 school year. The study matched the 12 comparison schools to the intervention schools based on ethnicity, socioeconomic status (measured by the proportion of students that participated in the free or reduced-price lunch program), and prior student mathematics scores (measured by the Iowa Test of Basic Skills [ITBS]). The analytic sample consisted of 732 students in 52 classes among the six intervention schools and 2,704 students among the 12 comparison schools (the number of classes in the comparison schools was not provided by the author).

Intervention group The six intervention schools used *Everyday Mathematics*® for the full 1998–99 school year. *Everyday Mathematics*® introduces mathematical concepts in a variety of ways during the school year and consists of daily lesson plans that usually begin with a Math Message, which is the focus of the lesson, and combines teacher-led discussions and hands-on group and individual activities during the class. In addition to the class components, students also keep a journal in which they write about mathematical concepts and work on homework assignments that are intended to reinforce practical experience with mathematics.

Comparison group The comparison group used the district's adopted textbook, *Mathematics in Action*, a traditional mathematics curriculum. *Mathematics in Action* focuses on the systematic understanding of concepts and algorithms in specific lesson plans with an emphasis on practice problems and repetition before new concepts are introduced.

Outcomes and measurement The primary outcome used to measure student mathematics achievement was the Total Math Score from the 1999 Texas Assessment of Academic Skills (TAAS), which was administered to students in April 1999. The study reports an overall score and three subtest scores on the TAAS that measure concepts, operations, and problem solving. In addition, for the overall score, the author reports subgroup results for students who were classified as Black, Hispanic, White, male, female, of low socioeconomic status, and other socioeconomic status.⁴ The mean national percentage ranking of student scores on the math section of the ITBS was used as a pretest. The ITBS pretest was administered to students in April 1998. For a more detailed description of the TAAS outcome measure, see Appendix B.

Support for implementation

Teachers and administrators in the intervention schools received 40 hours of initial training on *Everyday Mathematics*®, as recommended by the publisher. Teachers also received the curriculum's Teacher's Resource Package, which includes a variety of materials that help teachers successfully implement the program, such as the Teacher's Manual and Lesson Guide, a teacher's Resource Book, instructions on creating home and school links, and a materials kit that contains manipulatives (e.g., dice, rulers) used in the lessons.

Appendix B: Outcome measures for the mathematics achievement domain

Mathematics achievement	
<i>Texas Assessment of Academic Skills (TAAS) Mathematics Scale Score</i>	The TAAS is a criterion-referenced assessment developed by the Texas Education Agency (TEA) from the state-mandated curriculum used to assess higher order thinking and problem-solving skills. TEA reports an internal consistency reliability range of .88 to .92 for the TAAS assessment. The TAAS is scored on the same scale in each grade level, with a score of 2,100 resulting in a rating of “proficient,” and a score of 2,500 resulting in rating of “highly proficient.” Only the mathematics scores from the April 1999 assessment were used in this study. The TAAS mathematics assessment includes three subtests that measure achievement in mathematical concepts, mathematical operations, and mathematical problem-solving (as cited in Waite, 2000).

Appendix C: Findings included in the rating for the mathematics achievement domain

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			p-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
Waite, 2000 ^a								
Texas Assessment of Academic Skills (TAAS): Total Math Scale Score	Grades 3–5	18 schools/ 3,436 students	63.64 (11.46)	59.80 (14.81)	3.85	0.27	+11	< .01
Domain average for mathematics achievement (Waite, 2000)						0.27	+11	Not statistically significant
Domain average for mathematics achievement across all studies						0.27	+11	na

Table Notes: For mean difference, effect size, and improvement index values reported in the table, a positive number favors the intervention group and a negative number favors the comparison group. The effect size is a standardized measure of the effect of an intervention on outcomes, representing the average change expected for all individuals who are given the intervention (measured in standard deviations of the outcome measure). The improvement index is an alternate presentation of the effect size, reflecting the change in an average individual's percentile rank that can be expected if the individual is given the intervention. The statistical significance of the study's domain average was determined by the WWC. Some statistics may not sum as expected due to rounding. na = not applicable.

^a For Waite (2000), the *p*-value presented here was reported in the original study. A correction for clustering was needed and resulted in a WWC-computed *p*-value of .30 for the TAAS Total Math Score; therefore, the WWC does not find the result to be statistically significant. The reported group means are based on an ordinary least squares (OLS) regression model that adjusted for the ITBS pretest, gender, race/ethnicity, and socioeconomic status. This study is characterized as having a substantively important positive effect because the mean effect reported is positive and not statistically significant, but is substantively important, and no effects are negative and statistically significant, accounting for clustering and multiple comparisons. For more information, please refer to the WWC Procedures and Standards Handbook (version 3.0), p. 26.

Appendix D: Description of supplemental findings for the mathematics achievement domain

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			p-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
Waite, 2000 ^a								
Texas Assessment of Academic Skills (TAAS) Math: Concepts	Grades 3–5	18 schools/ 3,436 students	17.51 (2.55)	16.75 (3.11)	0.76	0.25	+10	< .01
TAAS Math: Operations	Grades 3–5	18 schools/ 3,436 students	13.08 (2.93)	12.20 (3.53)	0.89	0.26	+10	< .01
TAAS Math: Problem solving	Grades 3–5	18 schools/ 3,436 students	9.73 (3.59)	8.63 (3.60)	1.10	0.31	+12	< .01

Table Notes: The supplemental findings presented in this table are additional findings from studies in this report that meet WWC design standards with or without reservations, but do not factor into the determination of the intervention rating. For mean difference, effect size, and improvement index values reported in the table, a positive number favors the intervention group and a negative number favors the comparison group. The effect size is a standardized measure of the effect of an intervention on outcomes, representing the average change expected for all individuals who are given the intervention (measured in standard deviations of the outcome measure). The improvement index is an alternate presentation of the effect size, reflecting the change in an average individual's percentile rank that can be expected if the individual is given the intervention. Some statistics may not sum as expected due to rounding.

^a For Waite (2000), the *p*-values presented here were reported in the original study. A correction for clustering was needed and resulted in WWC-computed *p*-values of .34, .32, and .25 for the Concepts, Operations, and Problem solving subtests, respectively; therefore, the WWC does not find the results to be statistically significant. A correction for multiple comparisons was needed for the three TAAS math subtests, which did not change the significance of the results. The reported group means are based on unadjusted means.

Endnotes

¹ The descriptive information for this program was obtained from publicly available sources, including the publisher's and developer's websites (www.mheonline.com and <http://everydaymath.uchicago.edu>, downloaded September 2014). The WWC requests publishers review the program description sections for accuracy from their perspective. The program description was provided to the publisher in September 2014, and the WWC incorporated feedback from the publisher. Further verification of the accuracy of the descriptive information for this program is beyond the scope of this review.

² The literature search reflects documents publicly available by December 2014. The previous intervention report was released in September 2010. This report has been updated to include reviews of 30 studies that were not reviewed in the previous report. Of the additional studies, 24 were not within the scope of the review protocol for the Primary Mathematics topic area, and six were within the scope of the review protocol but did not meet WWC group design standards. A complete list and disposition of all studies reviewed are provided in the references. The report includes reviews of all previous studies that met WWC group design standards with or without reservations. The studies in this report were reviewed using the Standards from the WWC Procedures and Standards Handbook (version 3.0), along with those described in the Primary Mathematics review protocol (version 3.1). The evidence presented in this report is based on available research. Findings and conclusions may change as new research becomes available.

³ For criteria used in the determination of the rating of effectiveness and extent of evidence, see the WWC Rating Criteria on p. 18. These improvement index numbers show the average and range of individual-level improvement indices for all findings across the studies.

⁴ The subgroup results and subscores are not presented in this report because the author did not demonstrate baseline equivalence for the analyses.

Recommended Citation

U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse. (2015, November). *Primary Mathematics intervention report: Everyday Mathematics®*. Retrieved from <http://whatworks.ed.gov>

WWC Rating Criteria

Criteria used to determine the rating of a study

Study rating	Criteria
Meets WWC group design standards without reservations	A study that provides strong evidence for an intervention's effectiveness, such as a well-implemented RCT.
Meets WWC group design standards with reservations	A study that provides weaker evidence for an intervention's effectiveness, such as a QED or an RCT with high attrition that has established equivalence of the analytic samples.

Criteria used to determine the rating of effectiveness for an intervention

Rating of effectiveness	Criteria
Positive effects	Two or more studies show statistically significant positive effects, at least one of which met WWC group design standards for a strong design, AND No studies show statistically significant or substantively important negative effects.
Potentially positive effects	At least one study shows a statistically significant or substantively important positive effect, AND No studies show a statistically significant or substantively important negative effect AND fewer or the same number of studies show indeterminate effects than show statistically significant or substantively important positive effects.
Mixed effects	At least one study shows a statistically significant or substantively important positive effect AND at least one study shows a statistically significant or substantively important negative effect, but no more such studies than the number showing a statistically significant or substantively important positive effect, OR At least one study shows a statistically significant or substantively important effect AND more studies show an indeterminate effect than show a statistically significant or substantively important effect.
Potentially negative effects	One study shows a statistically significant or substantively important negative effect and no studies show a statistically significant or substantively important positive effect, OR Two or more studies show statistically significant or substantively important negative effects, at least one study shows a statistically significant or substantively important positive effect, and more studies show statistically significant or substantively important negative effects than show statistically significant or substantively important positive effects.
Negative effects	Two or more studies show statistically significant negative effects, at least one of which met WWC group design standards for a strong design, AND No studies show statistically significant or substantively important positive effects.
No discernible effects	None of the studies shows a statistically significant or substantively important effect, either positive or negative.

Criteria used to determine the extent of evidence for an intervention

Extent of evidence	Criteria
Medium to large	The domain includes more than one study, AND The domain includes more than one school, AND The domain findings are based on a total sample size of at least 350 students, OR, assuming 25 students in a class, a total of at least 14 classrooms across studies.
Small	The domain includes only one study, OR The domain includes only one school, OR The domain findings are based on a total sample size of fewer than 350 students, AND, assuming 25 students in a class, a total of fewer than 14 classrooms across studies.

Glossary of Terms

Attrition	Attrition occurs when an outcome variable is not available for all participants initially assigned to the intervention and comparison groups. The WWC considers the total attrition rate and the difference in attrition rates across groups within a study.
Clustering adjustment	If intervention assignment is made at a cluster level and the analysis is conducted at the student level, the WWC will adjust the statistical significance to account for this mismatch, if necessary.
Confounding factor	A confounding factor is a component of a study that is completely aligned with one of the study conditions, making it impossible to separate how much of the observed effect was due to the intervention and how much was due to the factor.
Design	The design of a study is the method by which intervention and comparison groups were assigned.
Domain	A domain is a group of closely related outcomes.
Effect size	The effect size is a measure of the magnitude of an effect. The WWC uses a standardized measure to facilitate comparisons across studies and outcomes.
Eligibility	A study is eligible for review and inclusion in this report if it falls within the scope of the review protocol and uses either an experimental or matched comparison group design.
Equivalence	A demonstration that the analysis sample groups are similar on observed characteristics defined in the review area protocol.
Extent of evidence	An indication of how much evidence supports the findings. The criteria for the extent of evidence levels are given in the WWC Rating Criteria on p. 18.
Improvement index	Along a percentile distribution of individuals, the improvement index represents the gain or loss of the average individual due to the intervention. As the average individual starts at the 50th percentile, the measure ranges from -50 to +50.
Intervention	An educational program, product, practice, or policy aimed at improving student outcomes.
Intervention report	A summary of the findings of the highest-quality research on a given program, product, practice, or policy in education. The WWC searches for all research studies on an intervention, reviews each against design standards, and summarizes the findings of those that meet WWC design standards.
Multiple comparison adjustment	When a study includes multiple outcomes or comparison groups, the WWC will adjust the statistical significance to account for the multiple comparisons, if necessary.
Quasi-experimental design (QED)	A quasi-experimental design (QED) is a research design in which study participants are assigned to intervention and comparison groups through a process that is not random.
Randomized controlled trial (RCT)	A randomized controlled trial (RCT) is an experiment in which eligible study participants are randomly assigned to intervention and comparison groups.
Rating of effectiveness	The WWC rates the effects of an intervention in each domain based on the quality of the research design and the magnitude, statistical significance, and consistency in findings. The criteria for the ratings of effectiveness are given in the WWC Rating Criteria on p. 18.
Single-case design	A research approach in which an outcome variable is measured repeatedly within and across different conditions that are defined by the presence or absence of an intervention.

Glossary of Terms

- Standard deviation** The standard deviation of a measure shows how much variation exists across observations in the sample. A low standard deviation indicates that the observations in the sample tend to be very close to the mean; a high standard deviation indicates that the observations in the sample tend to be spread out over a large range of values.
- Statistical significance** Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups. The WWC labels a finding statistically significant if the likelihood that the difference is due to chance is less than 5% ($p < .05$).
- Substantively important** A substantively important finding is one that has an effect size of 0.25 or greater, regardless of statistical significance.
- Systematic review** A review of existing literature on a topic that is identified and reviewed using explicit methods. A WWC systematic review has five steps: 1) developing a review protocol; 2) searching the literature; 3) reviewing studies, including screening studies for eligibility, reviewing the methodological quality of each study, and reporting on high quality studies and their findings; 4) combining findings within and across studies; and, 5) summarizing the review.

Please see the WWC Procedures and Standards Handbook (version 3.0) for additional details.



Practice
Guide



Quick
Review



Single Study
Review

An **intervention report** summarizes the findings of high-quality research on a given program, practice, or policy in education. The WWC searches for all research studies on an intervention, reviews each against evidence standards, and summarizes the findings of those that meet standards.

This intervention report was prepared for the WWC by Mathematica Policy Research under contract ED-IES-13-C-0010.